Published online 15 November 2014 *Nucleic Acids Research, 2015, Vol. 43, Database issue* ***D193–D196***
doi: 10.1093/nar/gku1173

LncRNA2Target: a database for differentially expressed genes after lncRNA knockdown or overexpression

Qinghua Jiang[1], Jixuan Wang[2], Xiaoliang Wu[3], Rui Ma[3], Tianjiao Zhang[3], Shuilin Jin[4], Zhijie Han[1], Renjie Tan[3], Jiajie Peng[3], Guiyou Liu[5], Yu Li[1] and Yadong Wang[3,*]

[1]School of Life Science and Technology, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China, [2]School of Software, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China, [3]School of Computer Science and Technology, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China, [4]Department of Mathematics, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China and [5]Genome Analysis Laboratory, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin 300308, China

Received August 09, 2014; Revised October 24, 2014; Accepted November 01, 2014

ABSTRACT

Long non-coding RNAs (lncRNAs) have emerged as critical regulators of genes at epigenetic, transcriptional and post-transcriptional levels, yet what genes are regulated by a specific lncRNA remains to be characterized. To assess the effects of the lncRNA on gene expression, an increasing number of researchers profiled the genome-wide or individual gene expression level change after knocking down or overexpressing the lncRNA. Herein, we describe a curated database named *LncRNA2Target*, which stores lncRNA-to-target genes and is publicly accessible at http://www.lncrna2target.org. A gene was considered as a target of a lncRNA if it is differentially expressed after the lncRNA knockdown or overexpression. *LncRNA2Target* provides a web interface through which its users can search for the targets of a particular lncRNA or for the lncRNAs that target a particular gene. Both search types are performed either by browsing a provided catalog of lncRNA names or by inserting lncRNA/target gene IDs/names in a search box.

INTRODUCTION

A large number of long non-coding RNAs (lncRNAs) have been identified (1), and emerging studies have revealed that lncRNAs are not transcriptional noise but play important roles in the regulation of a wide range of processes (2,3) and diseases (4–7). To help researchers to better understand lncRNAs, many important lncRNA databases have been developed. For example, Chen *et al.* developed a beneficial database named 'LncRNADisease' that store experimentally verified lncRNA–disease associations (6); Amaral *et al.* developed a wonderful database named 'lncRNAdb' that contains a comprehensive list of lncRNAs that have been shown to have, or to be associated with, biological functions in eukaryotes (8); Volders *et al.* constructed a helpful database titled 'LNCipedia' for the annotation of human lncRNA transcript sequences and structures (9); Yang *et al.* described a valuable database 'ChIPBase' to facilitate the comprehensive annotation and discovery of transcription factor binding maps and transcriptional regulatory relationships of lncRNAs from ChIP-Seq data.

Many studies have demonstrated that lncRNAs play crucial roles in the regulation of gene expression at epigenetic (10), transcriptional (11) and post-transcriptional (12) level. Recently, using low-throughput experimental technologies (e.g. quantitative reverse transcriptase-polymerase chain reaction (qRT-PCR) and western blot) and high-throughput experimental technologies (e.g. microarray or RNA-seq), an increasing number of researchers have attempted to profile the individual or genome-wide gene expression level change after knocking down or overexpressing a lncRNA of interest, and identified the differentially expressed genes as the target genes of the lncRNA for further functional analysis of the lncRNA. However, detailed information on these lncRNA-target relationships are scattered in relevant literature and there is no online repository that stores the information.

To fill this gap, we considered target genes of a lncRNA as the differentially expressed genes after knocking down or overexpressing the lncRNA, and developed a database named *LncRNA2Target* to manage all experimentally ver-

*To whom correspondence should be addressed. Tel: +86 451 86413316; Fax: +86 451 86403260; Email: ydwang@hit.edu.cn

Disclaimer: In our database, all differentially expressed genes after lncRNA perturbation were considered as targets of the lncRNA. Although this indeed suggests downstream regulation, it is no real proof of direct interaction. The effects on expression may be secondary or tertiary.

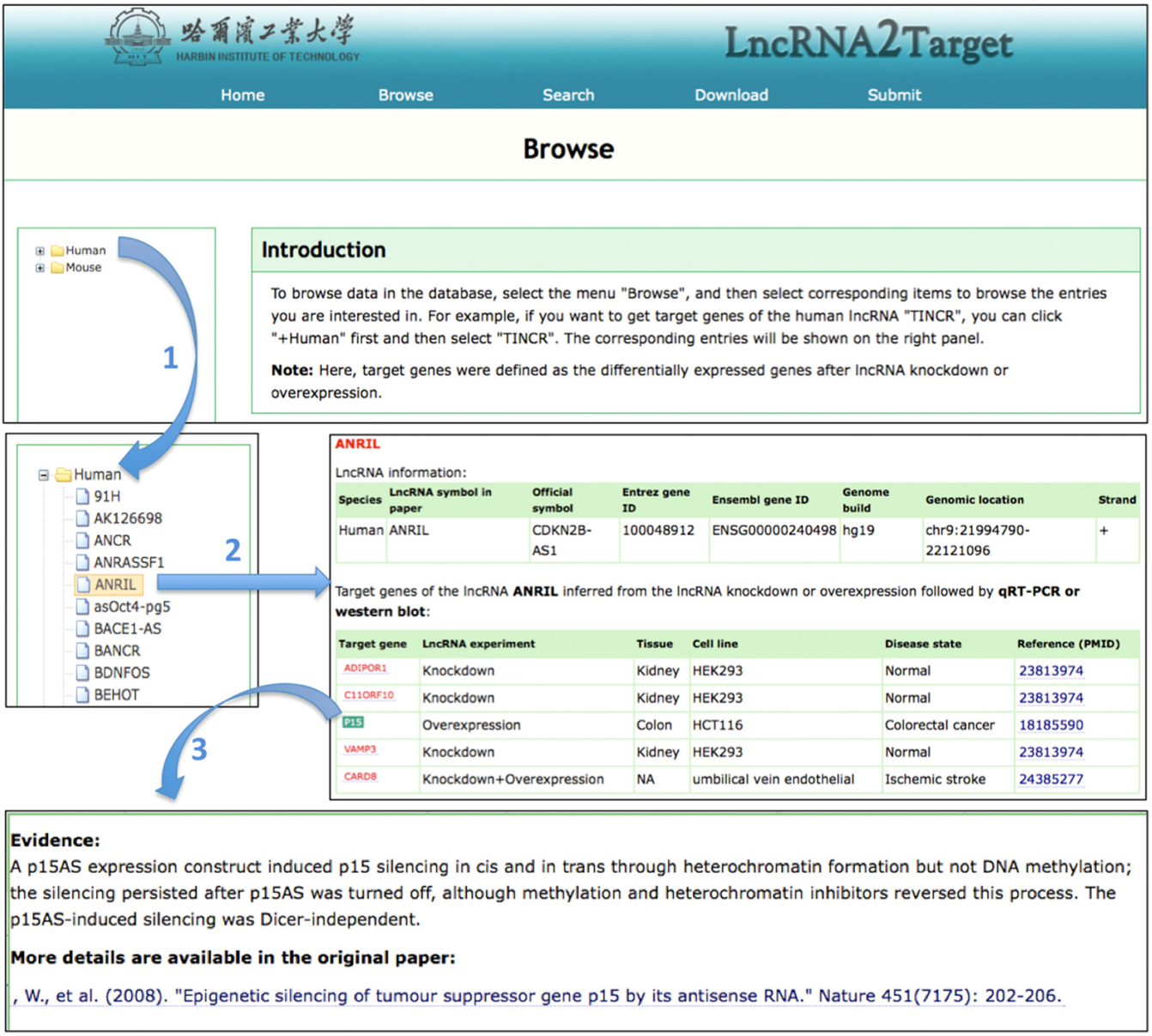

Figure 1. Screenshot of how to browse the target genes of lncRNA 'INRIL'.

Table 1. LncRNA knockdown or overexpression experiments followed by qRT-PCR or western blot to identify differentially expressed target genes

	No. of lncRNAs	No. of target genes	No. of lncRNA-target associations
Human	68	216	278
Mouse	26	95	118

Table 2. LncRNA knockdown or overexpression experiments followed by microarray or RNA-seq to identify differentially expressed target genes

	No. of lncRNAs	No. of target genes	No. of lncRNA-target associations
Human	14	11 389	26 133
Mouse	109	14 667	67 034

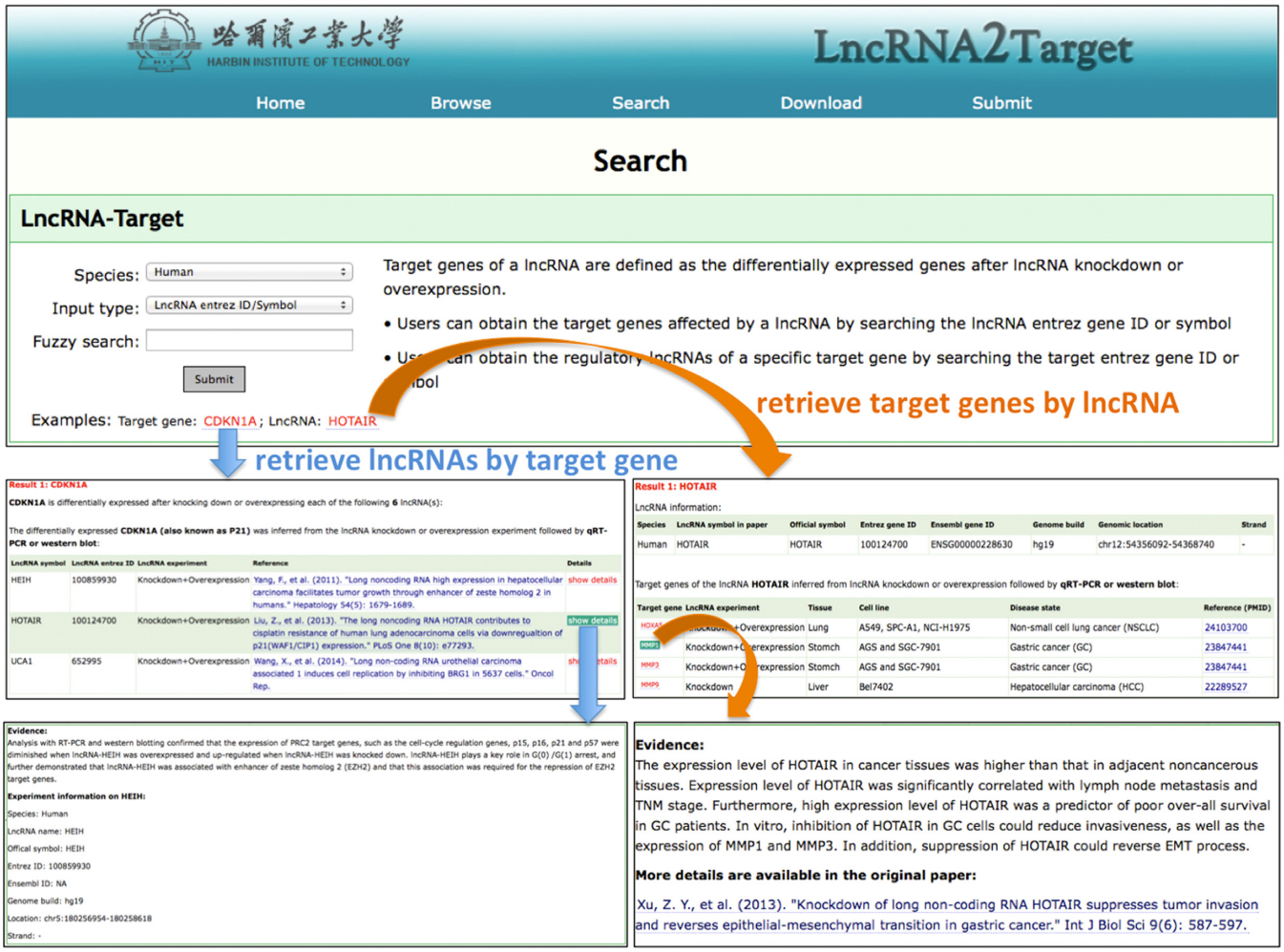

Figure 2. Screenshot of how to search target genes of lncRNA of interest or all lncRNAs that regulate a specific target gene.

ified lncRNA-target associations that were curated from published papers and microarray and RNA-seq data sets. Our database not only facilitates computational investigators to perform integrative analysis of the publicly available lncRNA targets but also enables experimental scientists to analyze their own data in the context of other related public data.

DATABASE OVERVIEW

The *LncRNA2Target* knowledge base aims to facilitate users to browse, search and download all literature-based lncRNA-target associations in the database, and submit novel data to the database. To extract the experimentally supported lncRNA–target associations, we first got all lncRNA papers published before 30 July 2014 by searching the PubMed literature database with keywords 'lncRNA', 'lincRNA', 'long non-coding RNA' or 'long intergenic non-coding RNA', respectively. Then, information on lncRNA–target associations was retrieved from the 150 papers and 11 high-throughput public data sets. Every association contains 18 items, such as species, lncRNA symbol/ID/location/strand, cell line in which the lncRNA knockdown or overexpression experiments was performed, disease state, target gene symbol/ID, supporting evidence for the lncRNA-target association and reference information. The detailed statistics on lncRNA-target associations are shown in Tables 1 and 2.

DATABASE ACCESS

User-friendly web interface was developed to facilitate users to browse, search and download the lncRNA-target association data, and upload new experientially verified lncRNA-target association to the database. Once approved by the submission review committee, the submitted record will be included in the database, and made available to the public in the coming release.

Browsing the database

Data in LncRNA2Target database can be browsed by lncRNA symbol. To browse the data, users first go into the 'Browse' page, and then select a species and a lncRNA they are interested in. For example, if the user wants to get target genes of the human lncRNA 'INRIL', you can click '+Human' first and then select 'INRIL'. The corresponding entries will be shown on the right panel. In this panel, basic

lncRNA information are shown, such as lncRNA symbol used in the original paper, official lncRNA symbol, its entrez or Ensembl gene ID, genomic location, strand, etc. In addition, the target genes of the lncRNA, cell line in which the lncRNA knockdown/overexpression experiments was performed, disease state and reference information are also shown in this panel. When users click the target symbol, a new window will be popped up, where the experimental evidence for supporting the lncRNA-target association is shown. An example on how to get the target genes of lncRNA 'INRIL' is shown in the Figure 1.

Searching the database

The LncRNA2Target database provides a 'Search' function for users to retrieve the lncRNA-target associations by lncRNA or target gene. The 'Search' is case-insensitive. Users can obtain the target genes affected by a lncRNA by searching the lncRNA entrez gene ID or symbol, and also can obtain the regulatory lncRNAs of a specific target gene by searching the target entrez gene ID or symbol. An example on how to retrieve the lncRNAs that regulate a specific gene named 'CDKN1A' and how to retrieve the target genes of lncRNA named 'HOTAIR' is shown Figure 2.

FUTURE DEVELOPMENT

As the number of lncRNA knockdown/overexpression experiments followed by qRT-PCR, western blot, microarray or RNA-seq to identify differential target genes increases exponentially, a submission page allows researchers to inform us about new publicly available data by inputting a PubMed ID or GEO accession. After manual curation and computational analysis, the new lncRNA target genes will be accessible in the coming release. In addition, we will search the PubMed and GEO database for extracting new available lncRNA-target associations, and update the LncRNA2Target regularly. With the joint effort between our lab and the lncRNA community, the LncRNA2Target will be a valuable resource for the lncRNA community.

FUNDING

Natural Science Foundation of China [61102149]; Fundamental Research Funds for the Central Universities [HIT NSRIF.2010057, HIT BRETIII.201219]; China National 863 High-Tech Program [2012AA02A602, 2012AA020404 and 2012AA02A604]. Funding for open access charge: China National 863 High-Tech Program [2012AA02A602].
Conflict of interest statement. None declared.

REFERENCES

1. Dunham,I., Kundaje,A., Aldred,S.F., Collins,P.J., Davis,C.A., Doyle,F., Epstein,C.B., Frietze,S., Harrow,J., Kaul,R. *et al.* (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57–74.
2. Ulitsky,I. and Bartel,D.P. (2013) lincRNAs: genomics, evolution, and mechanisms. *Cell*, **154**, 26–46.
3. Yuan,J.H., Yang,F., Wang,F., Ma,J.Z., Guo,Y.J., Tao,Q.F., Liu,F., Pan,W., Wang,T.T., Zhou,C.C. *et al.* (2014) A long noncoding RNA activated by TGF-beta promotes the invasion-metastasis cascade in hepatocellular carcinoma. *Cancer Cell*, **25**, 666–681.
4. Gutschner,T. and Diederichs,S. (2012) The hallmarks of cancer: a long non-coding RNA point of view. *RNA Biol.*, **9**, 703–719.
5. Spizzo,R., Almeida,M.I., Colombatti,A. and Calin,G.A. (2012) Long non-coding RNAs and cancer: a new frontier of translational research? *Oncogene*, **31**, 4577–4587.
6. Chen,G., Wang,Z., Wang,D., Qiu,C., Liu,M., Chen,X., Zhang,Q., Yan,G. and Cui,Q. (2013) LncRNADisease: a database for long-non-coding RNA-associated diseases. *Nucleic Acids Res.*, **41**, D983–D986.
7. Barnhill,L.M., Williams,R.T., Cohen,O., Kim,Y., Batova,A., Mielke,J.A., Messer,K., Pu,M., Bao,L., Yu,A.L. *et al.* (2014) High expression of CAI2, a 9p21-embedded long noncoding RNA, contributes to advanced-stage neuroblastoma. *Cancer Res.*, **74**, 3753–3763.
8. Amaral,P.P., Clark,M.B., Gascoigne,D.K., Dinger,M.E. and Mattick,J.S. (2011) lncRNAdb: a reference database for long noncoding RNAs. *Nucleic Acids Res.*, **39**, D146–D151.
9. Volders,P.J., Helsens,K., Wang,X., Menten,B., Martens,L., Gevaert,K., Vandesompele,J. and Mestdagh,P. (2013) LNCipedia: a database for annotated human lncRNA transcript sequences and structures. *Nucleic Acids Res.*, **41**, D246–D251.
10. Sanchez-Elsner,T., Gou,D., Kremmer,E. and Sauer,F. (2006) Noncoding RNAs of trithorax response elements recruit Drosophila Ash1 to Ultrabithorax. *Science*, **311**, 1118–1123.
11. Martens,J.A., Laprade,L. and Winston,F. (2004) Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. *Nature*, **429**, 571–574.
12. Carrieri,C., Cimatti,L., Biagioli,M., Beugnet,A., Zucchelli,S., Fedele,S., Pesce,E., Ferrer,I., Collavin,L., Santoro,C. *et al.* (2012) Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature*, **491**, 454–457.